(12) UK Patent Application (19) GB (11) 2 341 715 (13) A

(43) Date of A Publication 22.03.2000

(21) Application No 9820213.8

(22) Date of Filing 17.09.1998

(51) INT CL[7]
G11B 20/18

(52) UK CL (Edition R)
G5R RB33 RGB

(56) Documents Cited
EP 0795812 A1 EP 0717357 A2 EP 0569313 A2
EP 0569236 A2 EP 0485110 A2 EP 0450801 A2
WO 93/18455 A1 WO 91/20076 A1 WO 91/14982 A1
US 5651132 A

(71) Applicant(s)
Springtek Limited
(Incorporated in the United Kingdom)
Unit 3 Ashbrook Mews, Westbrook Street,
BLEWBURY, Oxon, OX11 9QA, United Kingdom

(72) Inventor(s)
Andrew Paul George Randall

(74) Agent and/or Address for Service
Atkinson Burrington
The Technology Park, 60 Shirland Lane, SHEFFIELD,
S9 3SP, United Kingdom

(58) Field of Search
UK CL (Edition P) G4A AES, G5R RAC RB33 RGB
INT CL[6] G06F 11/10, G11B 20/18
EDOC WPI

(54) Abstract Title
Magnetic disk redundant array

(57) A plurality of magnetic disk drives (301, 302, 303) are configured to store machine readable data in a protected way such that data is recoverable in the event of a single disk failure. The array of disks is housed for application directly into an existing disk bay of a computer (101). The array is connectable to the computer as if it were a single conventional computer disk and the drives are controlled by an operating system on the computer as if they were a single storage volume.

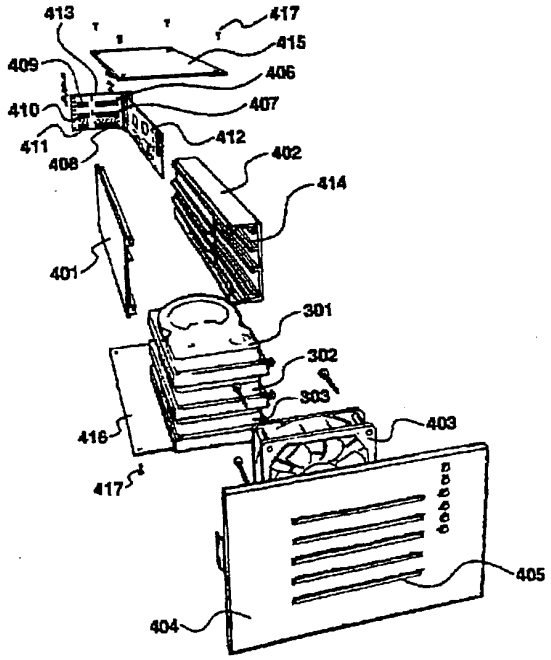

Figure 4

At least one drawing originally filed was informal and the print reproduced here is taken from a later filed formal copy.

This print takes account of replacement documents submitted after the date of filing to enable the application to comply with the formal requirements of the Patents Rules 1995

GB 2 341 715 A

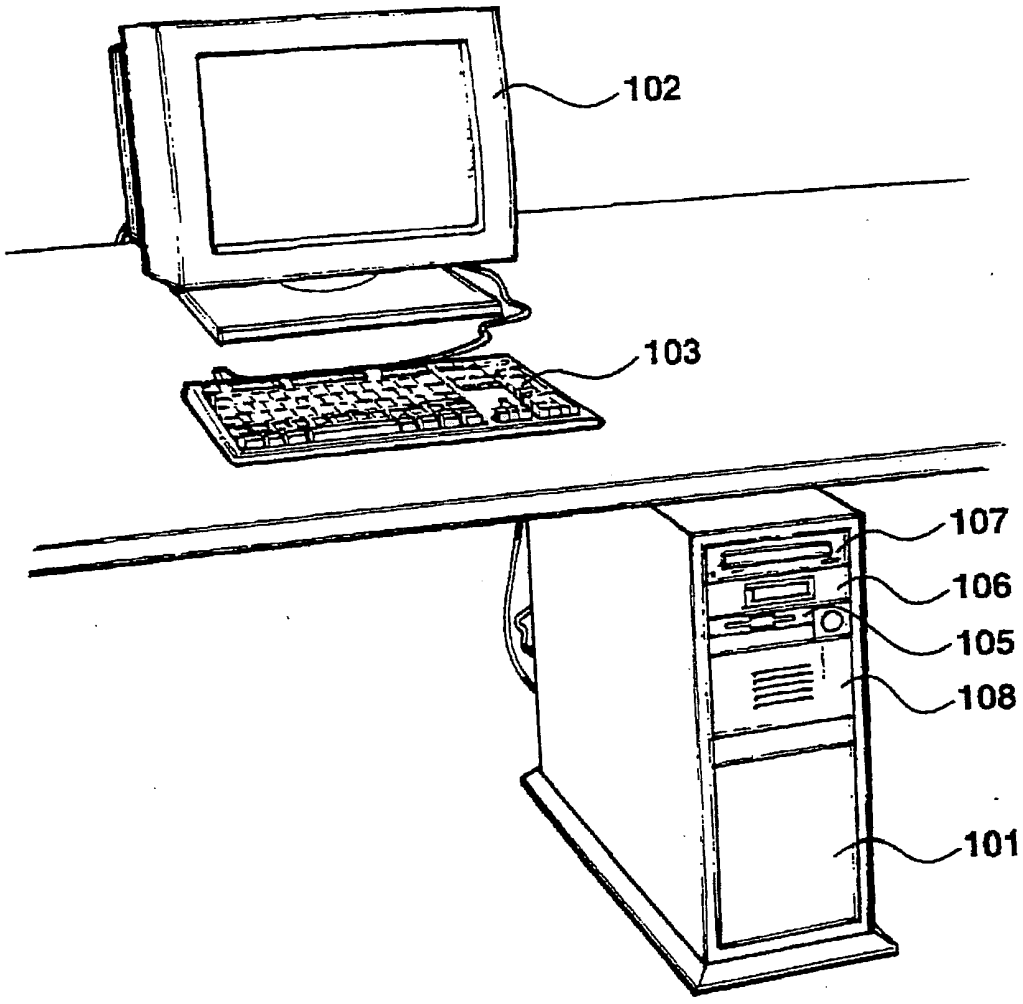

Figure 1

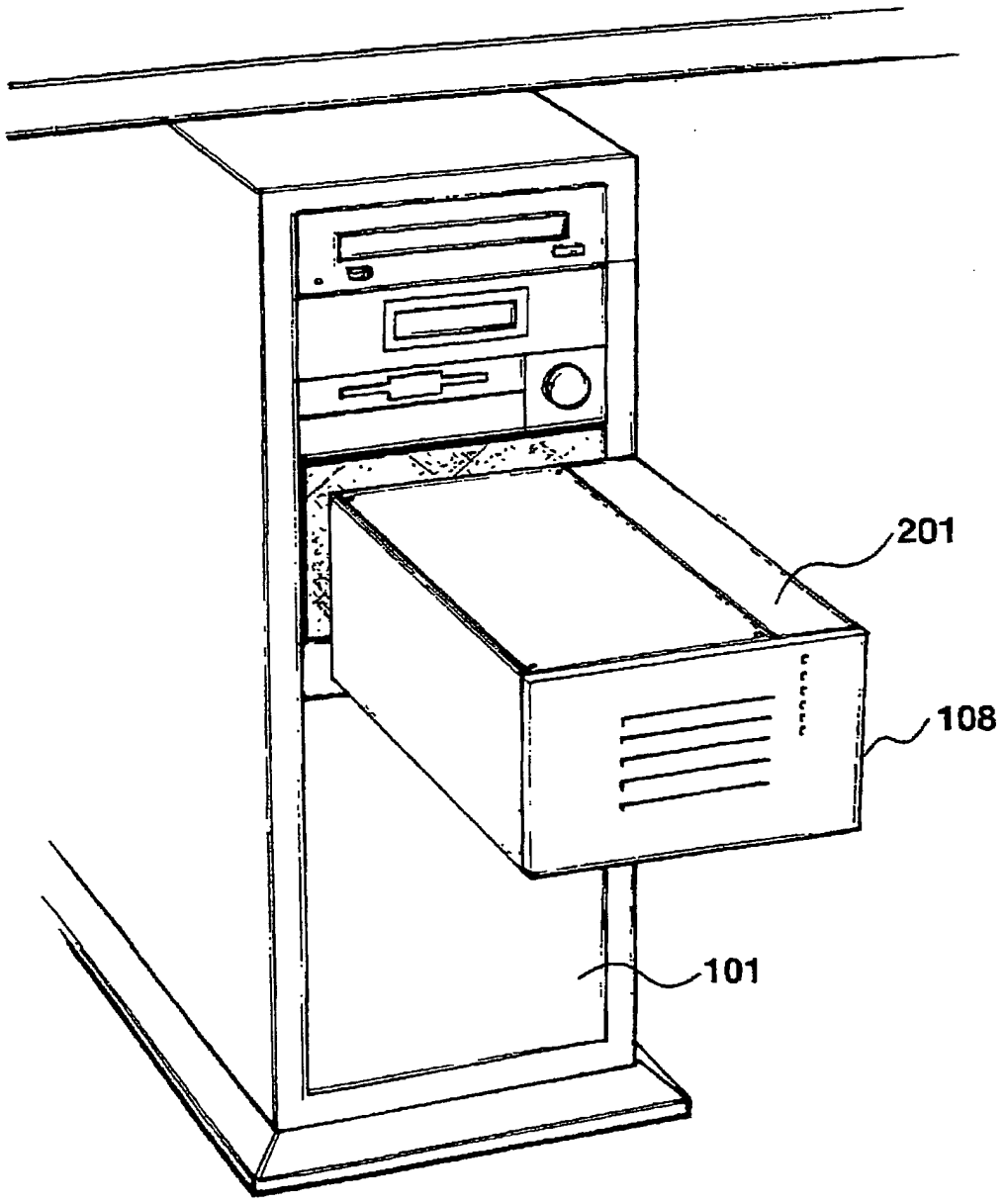

Figure 2

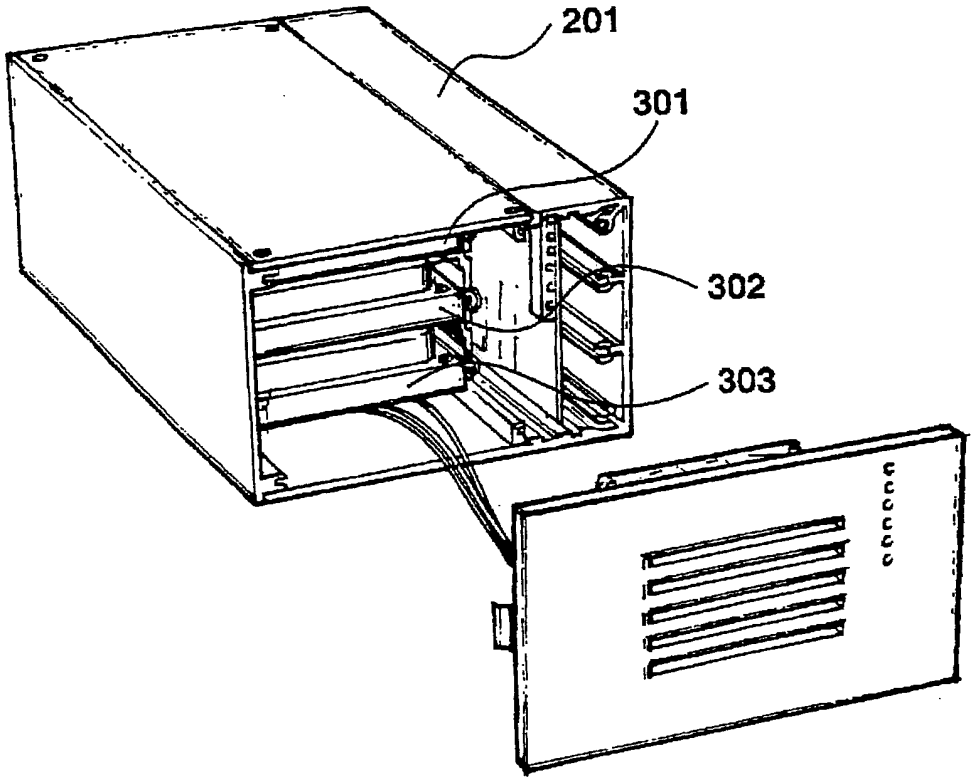

Figure 3

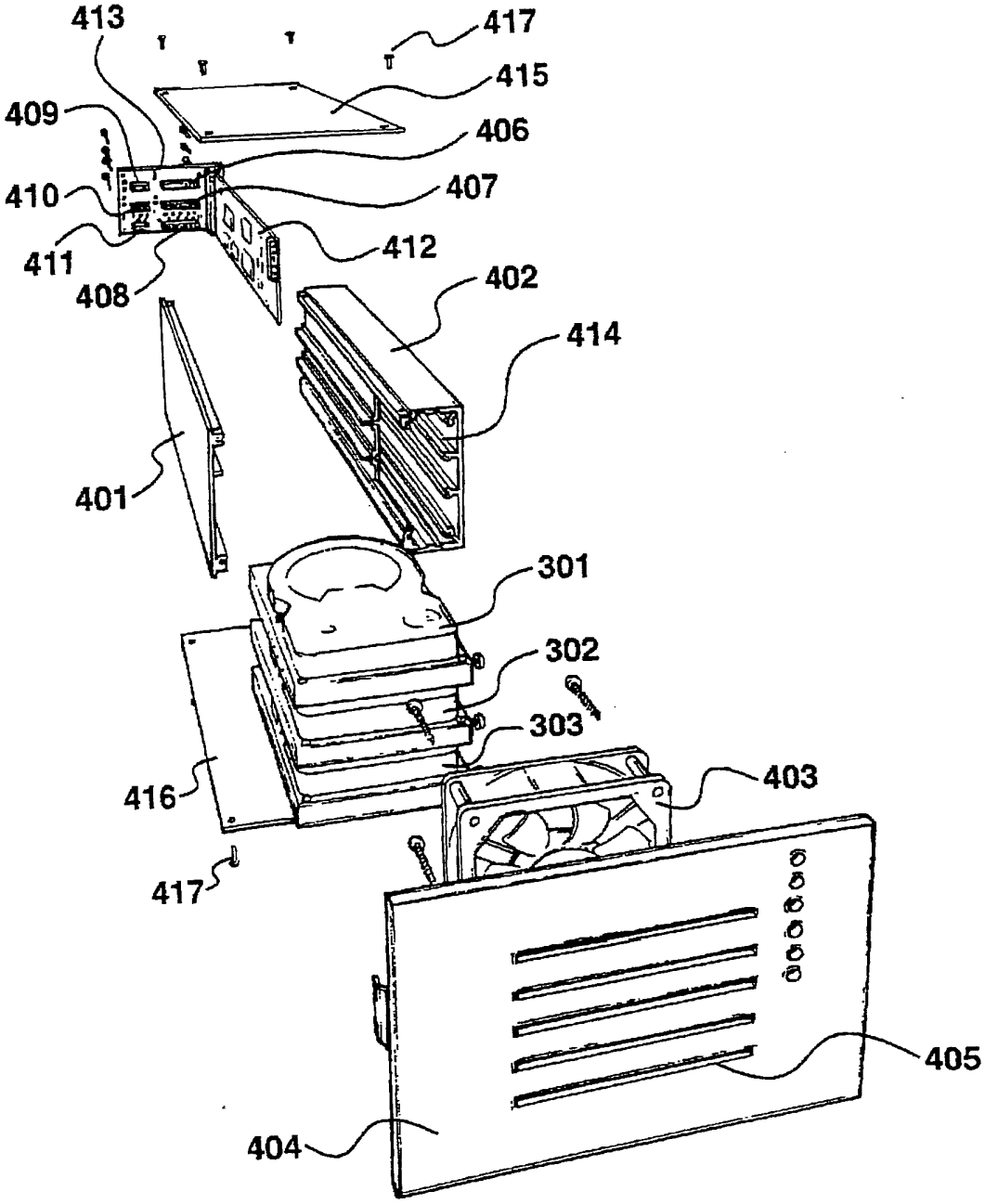
413
417
415
409
406
410
407
411
412
408
402
414
401
301
302
303
416
403
417
405
404

Figure 4

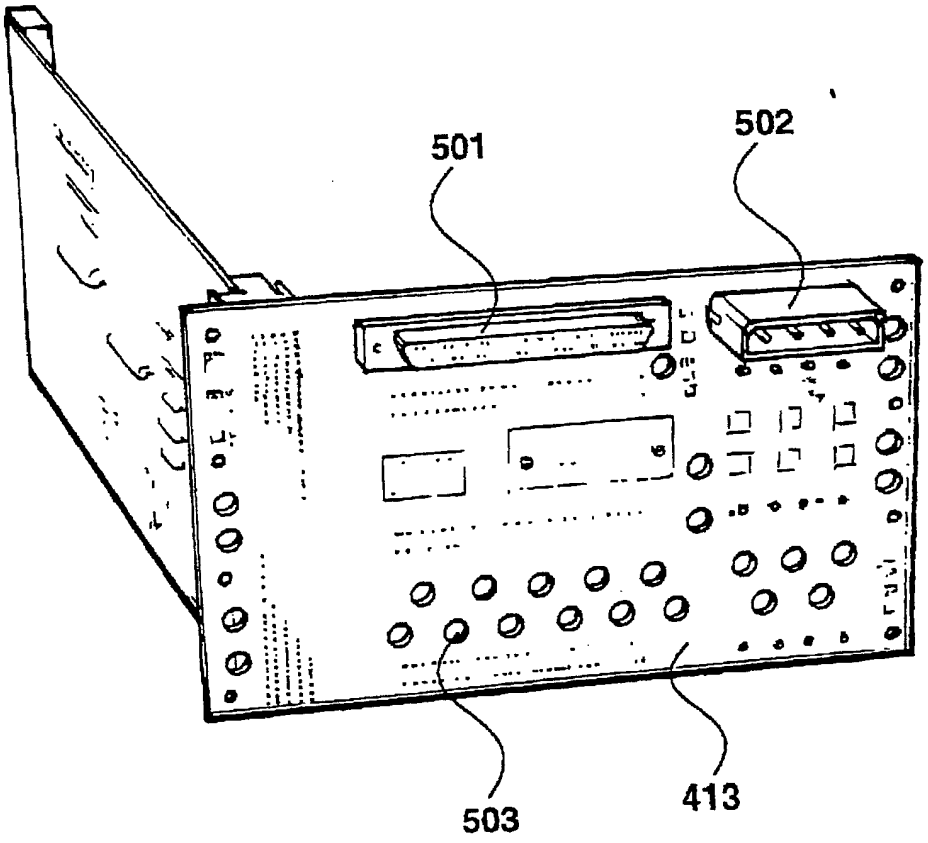

Figure 5

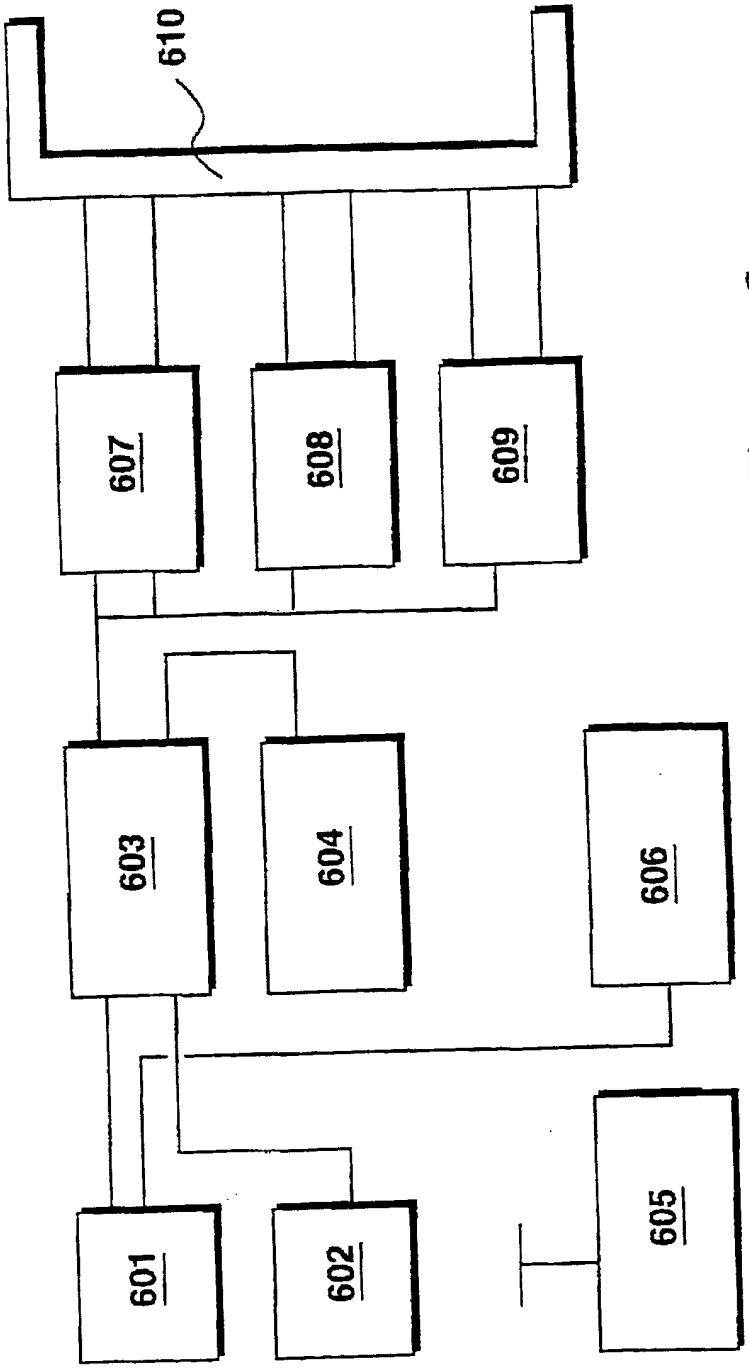

Figure 6

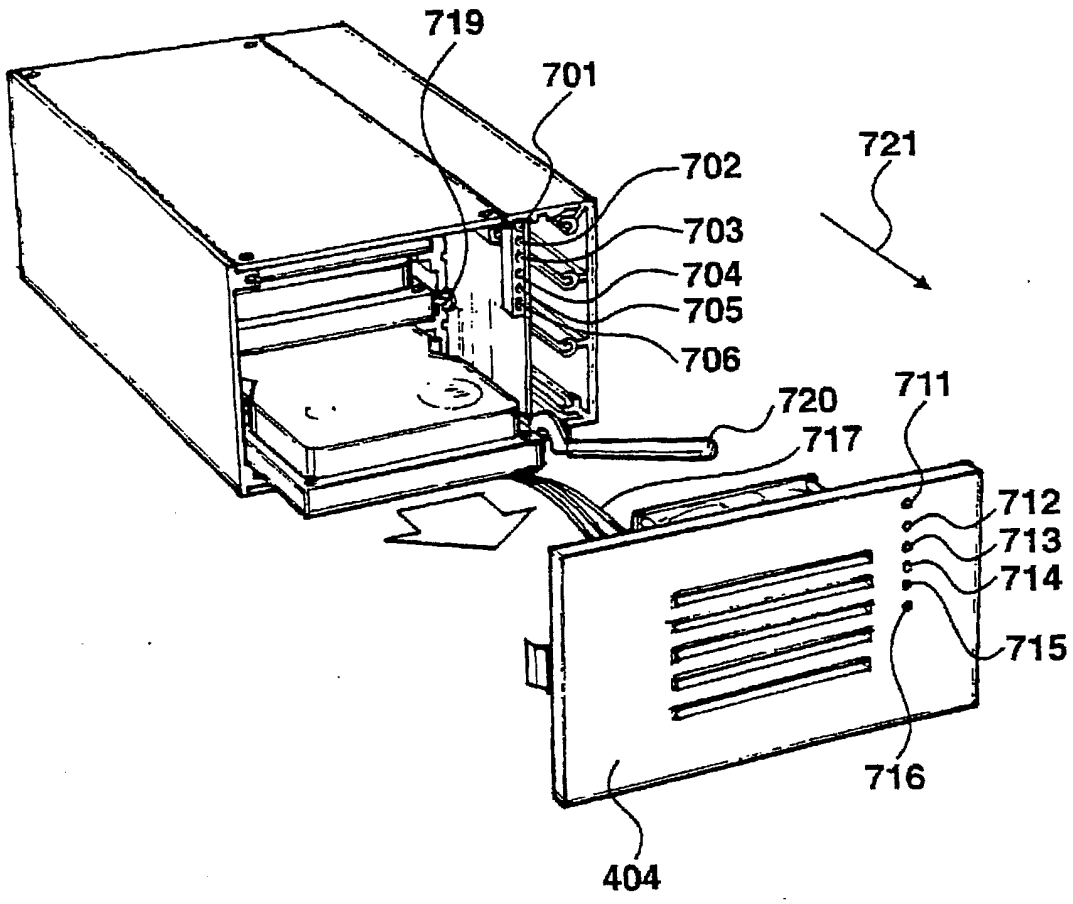
719
701
702
721
703
704
705
706
720
711
717
712
713
714
715
716
404

Figure 7

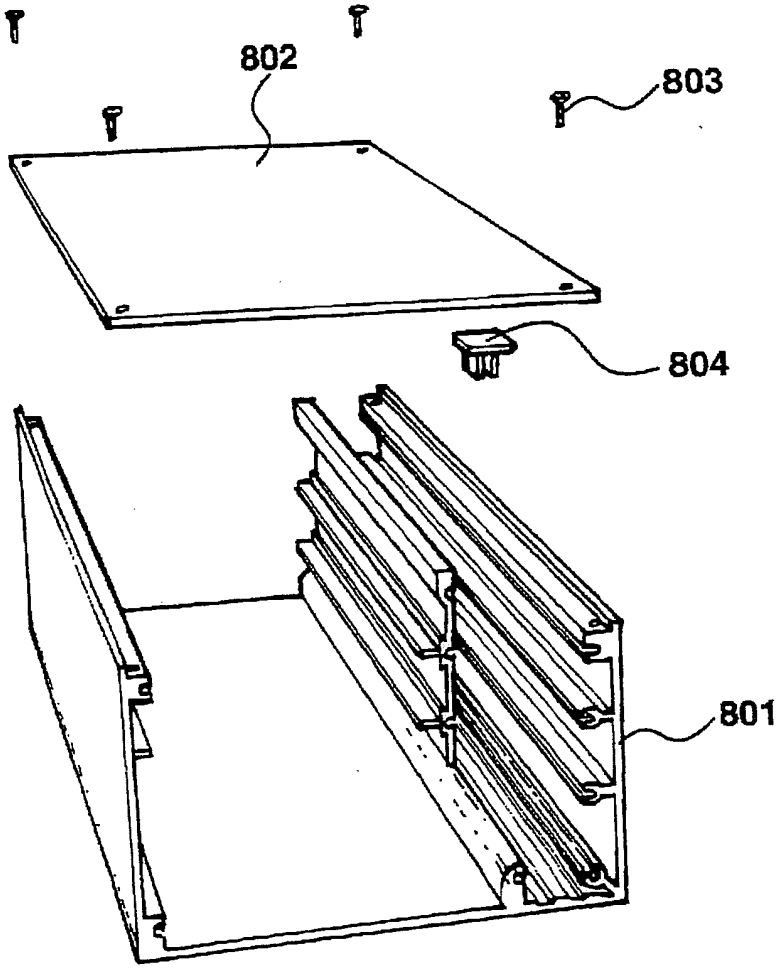

Figure 8

Data Storage

The present invention relates to an array of magnetic disks configured to store machine readable data in a protected way, such that data is recoverable in the event of disk failure.

Arrays configured to store machine readable data in a protected way are known and are often referred to as a redundant array of inexpensive disks, usually abbreviated to the acronym "RAID". Several RAID procedures are known and most of these share the approach of generating redundant data by an exclusive ORing process from which, in the event of any of the disks failing, all of the data can be reconstituted from the remaining operational disks.

When all of the disks are operational, the array is said to be working in its protected mode. In the event of one disk failure, the system may still remain operational, in that data may be read from the disks, but the data ceases to be protected and a further disk failure would result in data loss. With a single disk failure the system is said to be working in an unprotected mode at which point an operator would be advised that disk replacement is required and that the lost data needs to be reconstituted. Thus, a disk would be physically removed, replaced and then the lost data would be reconstituted on to the new disk.

As personal computer systems and workstations become more powerful, allowing more sophisticated software applications to be executed and the degree of data storage available in such systems increases, with disks containing several gigabytes of data now becoming widely used, a greater demand has been created for the installation of protected systems using disk redundancy. Complete RAID subsystems may be purchased for external connection but a problem with such known systems is that the cost can be very prohibitive. In many situations, the cost of such a RAID system

tends to be higher than the cost of a personal computer system. Thus, there is a requirement for providing RAID protection at reduced cost.

Personal computer systems are usually housed in desktop units or tower units having spare bays allowing additional disks to be received. Thus, it is possible for many hard disk drives to be included within a tower housing and additional interface cards may be provided if required. Thereafter, it is possible for the RAID calculations to be effected by the resident host CPU, such that the additional extra cost is quite modest. However, a major problem with such a configuration is that a significant processor overhead is required in order to perform the RAID calculations, resulting in a severe degradation in overall system performance.

According to a first aspect of the present invention, there is provided a plurality of data storage devices configured to store machine readable data in a protected way such that data is recoverable in the event of a single device failure, wherein the devices are housed for application directly into an existing disk bay for a computer; the devices are connectable to a disk interface as if they were a single conventional storage volume; and said devices are controlled by an operating system installed on a computer as if they were a single storage volume.

In a preferred embodiment, the disks are interfaced to an IDE connection and three disks may be received in respective IDE connections.

Preferably, the array presents a SCSI interface to a host computer and the array may be configured to be housed in two or more five and one quarter inch drive bays.

According to a second aspect of the present invention, there is provided a method of equipping a personal computer with a plurality of data storage devices configured as a redundant array by interfacing said devices to conventional five and one quarter inch drive bays, such that protected machine readable data is recoverable in the event of a single disk failure, comprising the steps of supporting the array within an existing disk bay for a

computer; connecting the array to the computer as if it were a single conventional computer disk; and controlling said drives by an operating system installed on a computer as if it were a single storage volume.

The invention will now be described by way of example only, with reference to the accompanying drawings, in which:

Figure 1 shows a personal computer system;

Figure 2 shows an array of disks being inserted into a computer system;

Figure 3 details the array shown in *Figure 2*;

Figure 4 shows an exploded view of the array identified in *Figure 3*;

Figure 5 shows a rear face view of the array back plane;

Figure 6 shows a circuit for implementing RAID calculations; and

Figure 7 illustrates the removal of a damaged disk from the array; and

Figure 8 shows an alternative embodiment for the extrusion identified in *Figure 4*.

A personal computer system is shown in *Figure 1* in which a main system tower **101** supplies visual information to a visual display unit **102** and receives manual commands via a keyboard **103**. The main system tower houses a central processing unit, memory circuits and other standard associated electronics as is well known in the art. The personal computer system may be an IBM PC type system, a Mackintosh system or any other computer type equipment used for individual use, possibly in a networked configuration. Alternatively, the main system tower **101** may constitute a network server, possibly running an appropriate server operating system, such as Windows NT server.

Tower **101** includes conventional five and one quarter inch disk bays. Within these disk bays a plurality of devices have been mounted, including a three and a half inch floppy disk drive **105**, a tape streamer **106**, a CD ROM drive **107** and an array of magnetic disks **108**, embodying the present invention.

Array **108** is detailed in *Figure 2* and is shown being installed into the main system tower **101**. The array **108** of magnetic disks is configured to store machine readable data in a protected way such that data is recoverable in the event of a single disk failure. The array of disks is housed for application directly into an existing disk bay of a computer, such as the main system tower **101**. The array is connectable to the computer as if it were a single conventional computer disk and the array is operated by an operating system installed on the computer as if it were a single disk.

Each empty drive bay is protected by a removable plastic cover and unit **107** locates within an aperture equivalent to the width of two bays, requiring the removal of two such covers. The array includes a housing **201**, locatable within the two bay aperture and towards its rear includes conventional power and data connectors; such that the housing as a whole is connected to the main system tower using a conventional SCSI connection. Thus, the main system perceives the disk array as if it were a single disk and the operating system, executed by the main system, controls the operation of the array using equivalent commands to those required for the operation of a single storage volume.

The array **107** is detailed in *Figure 3* and contains a total of three IDE drives **301**, **302** and **303**. An exploded view of the array is illustrated in *Figure 4*, which shows each of the individual IDE drives **301**, **302** and **303** being supported by aluminium extrusions, in the form of a left extrusion **401** and a right extrusion **402**. These extrusions hold the disk drives **301**, **302** and **303** firmly in place and facilitate the removal and replacement of individual disk drives when disk failure occurs.

Disk drives **301**, **302** and **303** are located in relatively close proximity and in order to maintain preferred operational temperatures, an electric fan **403** is positioned between the front of the disk drives and a front housing **404**. In this respect, the main front housing includes ventilation grilles **405**.

Each IDE drive **301**, **302** and **303** locates within a conventional IDE socket **406**, **407**, **408**, in addition to respective power supply sockets **409**, **410**, **411**. Thus, from the perspective of each IDE drive, the physical drives are located into sockets substantially similar to those found on an IDE bus of a standard computer system.

RAID calculations are performed within the device itself, using conventional hardware RAID circuitry mounted on circuit board **412**, having electrical connections to the back plane circuit board **413**. Right extrusion **402** defines a cavity **414**, configured to receive circuit board **412**. The extrusions **401** and **402** are held in position by an upper plate **415** and a lower plate **416**, secured by appropriate bolts **417**.

The rear face of back plane **413** is illustrated in *Figure 5*. The back plane includes a conventional SCSI socket **501** and a power supply socket **502**. The array therefore presents itself to the main system as a single disk drive, requiring a single disk drive connection via SCSI interface **501**.

Back plane **413** also includes rows of holes **503** to facilitate ventilation of the disks. Thus, cooling air is brought in through ventilation holes **405**, blown between the disks **301**, **302** and **303** and then exits through holes **503**.

The circuit implemented on board **412** is illustrated diagramatically in *Figure 6*. The circuit includes a central processing unit **601** which communicates with an input/output circuit **602** via a CPU bridge **603**. In addition, operation of CPU **601** is controlled by a CPU mode select circuit **604**. Power from the housing is directed to a three volt supply regulating circuit **605**, arranged to supply power to operational circuits via supply rails.

The CPU **601** receives data relating to the operational environment from an environmental detecting circuit **606**. This information may be received directly, as shown in *Figure 6*, or it may be directed via other control circuitry to allow combined environmental information to be returned to the CPU **601**.

Further output circuitry includes IDE controllers **607** and **608** and a SCSI controller **609**. These circuits communicate with the back plane sockets via a one hundred and eighty way connector **610**.

Input/output circuit **602** supplies driving current to six LED's **701**, **702**, **703**, **704**, **705** and **706** shown in *Figure* 7. Each of these LED's is visible by means of respective holes **711**, **712**, **713**, **714**, **715** and **716** in the front panel **404**. Each LED is a Hewlett Packard HSMF-C655 and actually includes a green LED and a red LED which may be operated independently.

LED **701** indicates the overall operational integrity of the system and primarily confirms that CPU **601** is operating correctly. Thus, when the system is fully operational, LED **701** is illuminated green. Alternatively, if faults have been detected within the controller, LED **701** is illuminated red.

LED **702** represents the environmental monitoring status and is primarily concerned with operational temperature. Environmental circuit **606** includes a temperature sensor and a fault condition is generated if this sensor detects that operational temperatures have become excessive. In addition, a tachometer is associated with fan **403** and a fault condition is generated if this detects that rotation of the fan has ceased. Malfunction of fan **403** represents a serious problem in that this could result in all three drives being permanently damaged such that no protection is offered by the RAID configuration. The system also detects the presence of appropriate voltages on voltage supply rails, as supplied by power supply unit **605** in addition to detecting appropriate terminator power on the SCSI bus.

When the supply rail voltages are correct, SCSI terminator power is correct, the fan is operational and the system is working at its optimal operational temperature, LED **702** is illuminated green. If the system encounters problems and diverges from its preferred operational characteristics, such a condition is detected and LED **702** is illuminated orange. Under these conditions, further operation of the system is permitted but warnings may be generated to the effect that a job should be closed

down and that the device should be investigated. If problems continue and the situation worsens, particularly if the operational temperature becomes very high, LED **702** is illuminated red. Under these conditions, power to the drives is removed and an error condition is generated such that further access to the drives is not permitted.

LED **703** indicates that the SCSI connection is fully operational by being illuminated green. Furthermore, when the SCSI bus is actually in use, LED **703** is illuminated orange.

LED's **704**, **705** and **706** represents operational characteristics of the individual drives **301**, **302** and **303** respectively. When the drives are operational, the LED's are illuminated green and then illuminated orange when the actual data transfer takes place. Furthermore, if a disk error is detected, to the effect that an individual disk has failed, its respective LED is illuminated red.

In response to a single disk failure, it is preferable for the system to be placed off-line and for the damaged disk to be replaced immediately so that the lost data may be reconstituted and the system returned to protected mode operation. In order to replace a disk, the front panel is removed, an operation facilitated by the front panel **404** being retained simply to the main housing by means of an interference connection. Having removed the front panel **404** it is restrained by wires **717** required for supplying electrical power to fan **403**.

The disk drives include tapped holes towards their front-right corner and each of said tapped holes receives a threaded stud **719**. Stud **719** allows its respective disk **301** to **303** to be removed by the application of a stud hook **720**. Force is applied in the direction of arrow **721**, thereby forcing the respective disk drive away from its IDE and data sockets, such as sockets **406** and **409** etc.

An alternative embodiment is illustrated in *Figure 8*. In this embodiment, side panels and a base panel are fabricated as a single

extrusion **801**. The housing is then completed by the application of a top panel **802**. The top panel **802** is secured to the lower extrusion **801** by means of bolts **803** and circuitry held within the extrusion is further secured by an adhesive clip **804**.

Claims

1. A plurality of data storage devices configured to store machine readable data in a protected way such that data is recoverable in the event of a single device failure, wherein

the devices are housed for application directly into an existing disk bay for a computer;

the devices are connectable to a disk interface as if they were a single conventional storage volume; and

said devices are controlled by an operating system installed on a computer as if they were a single storage volume.

2. Data storage devices according to claim 1, wherein said storage devices are magnetic disk drives.

3. Data storage devices according to claim 2, wherein the magnetic disks are interfaced to an IDE connection.

4. Data storage devices according to claim 3, wherein three disks are received in respective IDE connections.

5. Data storage devices according to any of claims 1 to 3, wherein said devices present a SCSI interface to a host computer.

6. Data storage device according to any of claims 1 to 5, configured to be housed in two or more five and one quarter inch drive bays.

7. Data storage devices according to any of claims 1 to 6, including means for detecting when said devices are operating in non-ideal conditions.

8. Data storage devices according to claim 7, including means for detecting when said devices are operating at excessive temperatures.

9. Data storage devices according to claim 7 or claim 8, including means for detecting non-operation of a cooling fan.

10. Data storage devices according to claim 7 or claim 8, including means for directly detecting an excessive operational temperature.

11. Data storage devices according to any of claims 7 to 10, including means for removing drive power to said devices upon detecting a non-ideal operating condition.

12. Data storage devices according to any of claims 1 to 11, including a detachable front panel and a cooling fan secured to said front panel, including ventilation openings arranged to direct a cooling air-stream between the individual devices.

13. A plurality of data storage devices according to any of claims 1 to 12, wherein said devices are connectable in a computer housing and the devices are controlled by the operating system of said computer.

14. A method of equipping a personal computer with a plurality of data storage devices configured as a redundant array by interfacing said devices to conventional five and one quarter inch drive bays, such that protected machine readable data is recoverable in the event of a single disk failure, comprising the steps of

supporting the array within an existing disk bay for a computer;

connecting the array to the computer as if it were a single conventional computer disk; and

controlling said drives by an operating system installed on a computer as if it were a single storage volume.

15. A method according to claim **14**, wherein said data storage devices are magnetic disk drives.

16. A method according to claim **15**, wherein said magnetic disk drives are interfaced to an IDE connection.

17. A method according to claim **16**, wherein three disks are received in respective IDE connections.

18. A method according to any of claims **14** to **17**, wherein said devices present a SCSI interface to a host computer.

19. A method according to any of claims **14** to **18**, wherein said devices are housed in two or more five and one quarter inch drive bays.

20. A method according to any of claims **14** to **19**, wherein non-ideal operating conditions for said devices are detected.

21. A plurality of data storage devices substantially as herein described with reference to the accompanying Figures.

22. A method of equipping a personal computer substantially as herein described with reference to the accompanying Figures.

The Patent Office

INVESTOR IN PEOPLE

Application No: GB 9820213.8 **Examiner:** Julyan Elbro
Claims searched: 1 to 22 **Date of search:** 4 January 1999

Patents Act 1977
Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.Q): G5R (RGB, RB33, RAC); G4A (AES)

Int Cl (Ed.6): G06F 11/10; G11B 20/18

Other: EDOC WPI

Documents considered to be relevant:

Category	Identity of document and relevant passage		Relevant to claims
X	EP 0795812 A1	HITACHI see figure 1 and pages 2-3.	1-20
X	EP 0717357 A2	SYMBIOSIS LOGIC see abstract and figure 2.	1-20
X	EP 0569313 A2	INTERNATIONAL BUSINESS MACHINES see abstract and figures 1 and 3.	1-20
X	EP 0569236 A2	COMPAQ see figure 2 and pages 2-4.	1-20
X	EP 0485110 A2	ARRAY TECHNOLOGY see abstract.	1-20
X	EP 0450801 A2	INTERNATIONAL BUSINESS MACHINES see abstract, column 22 line 34 to column 23 line 11, and column 27 lines 15-25.	1-20
X	WO 93/18455 A1	ARRAY TECHNOLOGY see abstract, figure 1, and page 10 lines 2-26.	1-20
X	WO 91/20076 A1	STORAGE TECHNOLOGY see abstract and figure 1.	1-20
X	WO 91/14982 A1	SF2 CORPORATION see abstract and figures 1 and 2.	1-20

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.

The Patent Office

13

INVESTOR IN PEOPLE

Application No: GB 9820213.8 **Examiner:** Julyan Elbro
Claims searched: 1 to 22 **Date of search:** 4 January 1999

Category	Identity of document and relevant passage	Relevant to claims
X	US 5651132 A HITACHI see abstract and figure 1.	1-20

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.